A New Methodological Approach for Studying Intergenerational Mobility with an Application to Swiss Data

Ben Jann and Simon Seiler

Institute of Sociology, University of Bern
ben.jann@soz.unibe.ch, simon.seiler@soz.unibe.ch

Conference on "(Persistent) Inequalities Revisited:
Social Origin, Education, and Social Mobility",
Monte Verità, Ascona, July 26–31, 2015

Overview

- Introduction
- Methods
  - Log-multiplicative layer effect model
  - The PRE approach
- Data
- Results
  - Comparison of LMLEM and PRE
  - Smoothed PRE
  - PRE with control variables
  - PRE with multiple origin variables
  - Direct and indirect origin effects
- Conclusions

Introduction

- As a methodology to analyze the development of social mobility based on categorical variables such as class or educational attainment, Erikson and Goldthorpe (1992) and Xie (1992) independently proposed a variant of the log-linear model known as the *uniform difference model* (Unidiff) or the *log-multiplicative layer effect model* (LMLEM).
- The model has been the standard tool in social mobility research since. The model, however, also has some limitations. We therefore propose an alternative approach.
- Our approach is based on the concept of proportional reduction of error (PRE). It quantifies the degree to which information about parents helps predicting the status of the children.
- The approach, we believe, is more flexible than the LMLEM and provides results that are easier to interpret.

Log-Multiplicative Layer Effect Model

- Starting point of the LMLEM is a simple two-way table of origin and destination, called a "mobility table," such as the following:

Parent's education	Respondent's education: compulsory or less	secondary vocational	secondary general	tertiary vocational	tertiary academic	Total
compulsory or less	201	332	16	71	68	690
secondary vocational	43	755	30	164	277	1269
secondary general	5	21	5	23	23	77
tertiary vocational	9	68	16	130	60	283
tertiary academic	16	78	13	41	317	466
Total	275	1254	81	429	745	2784

Source: see data section. Selection: males, birth cohorts 1969-82

Log-Multiplicative Layer Effect Model

- Such a two-dimensional mobility table can be formalized as follows, where F_{ij} are observed cell counts and $F_{i.}$ and $F_{.j}$ are row and column totals.

	1	...	j	...	J	Total
1	F_{11}	...	F_{1j}	...	F_{1J}	$F_{1.}$
$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
i	F_{i1}	...	F_{ij}	...	F_{iJ}	$F_{i.}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
I	F_{I1}	...	F_{Ij}	...	F_{IJ}	$F_{I.}$
Total	$F_{.1}$	...	$F_{.j}$	...	$F_{.J}$	$F_{..}$

- In a log-linear model, the cell counts are expressed as a multiplicative function:

$$F_{ij} = \tau_{..} \cdot \tau_{i.} \cdot \tau_{.j} \cdot \tau_{ij}, \quad i = 1, \ldots, I, \; j = 1, \ldots, J$$

Log-Multiplicative Layer Effect Model

- Now think of a table with an additional dimension (e.g. time points or birth cohorts).
- The saturated log-linear model for such a three-dimensional table is

$$F_{ijk} = \tau_{...} \cdot \tau_{i..} \cdot \tau_{.j.} \cdot \tau_{..k} \cdot \tau_{i.k} \cdot \tau_{.jk} \cdot \tau_{ij.} \cdot \tau_{ijk}$$

with $i = 1, \dots, I$, $j = 1, \dots, J$, and $k = 1, \dots, K$

	1	$\dots$	j	$\dots$	J	Total
1	F_{111}	$\dots$	F_{1j1}	$\dots$	F_{1J1}	$F_{1.1}$
$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
i	F_{i11}	$\dots$	F_{ij1}	$\dots$	F_{iJ1}	$F_{i.1}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
I	F_{I11}	$\dots$	F_{Ij1}	$\dots$	F_{IJ1}	$F_{I.1}$
Total	$F_{.11}$	$\dots$	$F_{.j1}$	$\dots$	$F_{.J1}$	$F_{..1}$

$\vdots$

	1	$\dots$	j	$\dots$	J	Total
1	F_{11k}	$\dots$	F_{1jk}	$\dots$	F_{1Jk}	$F_{1.k}$
$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
i	F_{i1k}	$\dots$	F_{ijk}	$\dots$	F_{iJk}	$F_{i.k}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
I	F_{I1k}	$\dots$	F_{Ijk}	$\dots$	F_{IJk}	$F_{I.k}$
Total	$F_{.1k}$	$\dots$	$F_{.jk}$	$\dots$	$F_{.Jk}$	$F_{..k}$

$\vdots$

	1	$\dots$	j	$\dots$	J	Total
1	F_{11K}	$\dots$	F_{1jK}	$\dots$	F_{1JK}	$F_{1.K}$
$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
i	F_{i1K}	$\dots$	F_{ijK}	$\dots$	F_{iJK}	$F_{i.K}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
I	F_{I1K}	$\dots$	F_{IjK}	$\dots$	F_{IJK}	$F_{I.K}$
Total	$F_{.1K}$	$\dots$	$F_{.jK}$	$\dots$	$F_{.JK}$	$F_{..K}$

Log-Multiplicative Layer Effect Model

- To ease interpretation, Xie (1992) proposed a simplified model in which $\tau_{ij.} \cdot \tau_{ijk}$ is replaced by $\exp(\psi_{ij.} \cdot \phi_{..k})$. This is called the log-multiplicative layer effect model:

$$F_{ijk} = \tau_{...} \cdot \tau_{i..} \cdot \tau_{.j.} \cdot \tau_{..k} \cdot \tau_{i.k} \cdot \tau_{.jk} \cdot \exp(\psi_{ij.} \cdot \phi_{..k})$$
$$i = 1, \ldots, I, \; j = 1, \ldots, J, \; k = 1, \ldots, K$$

- The $\psi_{ij.}$ parameters capture the overall pattern of dependencies between origin and destination.
- The $\phi_{..k}$ are cohort-specific scaling factors. That is, the higher $\phi_{..k}$, the more pronounced is the pattern of dependencies in cohort k and, hence, the stronger is the association between origin and destination, assuming that there is a stable basic pattern of associations across cohorts.
- To identify the model, constraints have to be placed on $\phi_{..k}$. Following Xie (1992), we use constraint $\sum_k \phi_{..k}^2 = 1$.

The PRE Approach

- The LMLEM provides a parsimonious and intuitive way to describe changes in social mobility across time and also allows testing against a null model with time-constant origin effects. However, the model has a number of limitations.
  - First, it assumes a common baseline pattern of associations that remains constant over time. This assumption may be violated so that results are biased.
  - Second, it is difficult to extend the model to include control variables.
  - Third, there is no clear interpretation of the absolute values of $\phi_{..k}$. In fact, the overall level of the $\phi_{..k}$ parameters is meaningless, because the sum over $\phi^2_{..k}$ is restricted to 1. This implies that $\phi_{..k}$ cannot be compared across models.
- We therefore propose an alternative approach that is based on (categorical) regression models and the PRE principle.

The PRE Approach

- General ideas:
  - The stronger the effect of the status of the parents on the status of their children, the lower is intergenerational mobility.
  - The „strength“ of an effect is easy to conceptualize for single regression coefficients. Things get more complicated, however, if we have to determine the strength of an effect that comprises multiple parameters.
  - Instead of thinking in terms of model parameters, however, we can ask how “useful” the information on parents is to predict the status of their children.
  - The better the position of children can be predicted based on parents characteristics, the stronger is the influence of social origin and the lower is social mobility.
  - To quantify the predictive power of parents’ characteristics we can resort to the statistical concept of the Proportional Reduction of Error (PRE).

The PRE Approach

- Formally:

$$PRE = \frac{E_0 - E_1}{E_0} = 1 - \frac{E_1}{E_0}$$

  where E_0 is the sum of prediction errors under limited information and E_1 is the sum of prediction errors under full information.

- Different error rules can be applied, yielding different PRE measures. Because our dependent variables are categorical, an entropy-based definition (see Theil 1970) appears appropriate:

$$E_m = -\sum_{i=1}^{N} w_i \ln\left(\widehat{p}_m(Y = y_i)\right), \quad m = 0, 1$$

  where w_i is the respondent's survey weight and $\widehat{p}_m(Y = y_i)$ is the predicted probability of the dependent variable taking on observed value y_i under model m.

The PRE Approach

- To estimate $\widehat{p}_m(Y = y_i)$ we use multinomial logit models.
- That is, the probabilities under restricted information are modeled as

$$p_0(Y = y_i) = \frac{\exp(\beta_{y_i} Z_i)}{\sum_{\ell=1}^{J} \exp(\beta_\ell Z_i)}$$

  where Z_i is a vector of control variables (possibly just a constant) and β_ℓ is an outcome-specific coefficient vector.
- Likewise, the probabilities under full information are modeled as

$$p_1(Y = y_i) = \frac{\exp(\beta_{y_i} Z_i + \gamma_{y_i} X_i)}{\sum_{\ell=1}^{J} \exp(\beta_\ell Z_i + \gamma_\ell X_i)}$$

  where X_i is a vector of parents' characteristics.

The PRE Approach

- To summarize, the procedure works as follows:
  - Divide the sample into a number of cohorts.
  - In each cohort, estimate two multinomial logit models, one without parent information and one including parent information.
  - Obtain the prediction error sums E_0 and E_1 from the two models and compute PRE.
  - Standard errors can be obtained by bootstrap.

Data

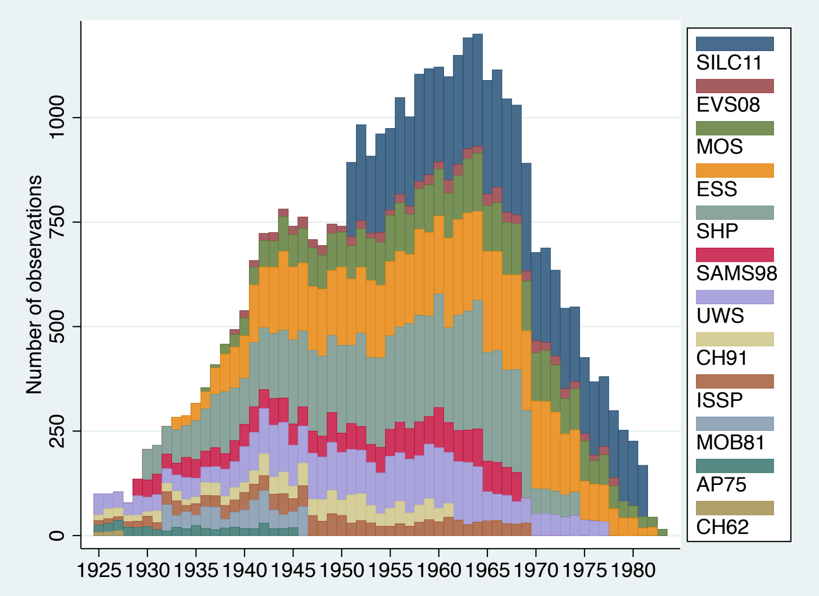
- Required are data that contain the relevant status variables for the respondents as well as information about education and occupation of parents.
- Most Swiss large-scale surveys, such as the official surveys by the Federal Statistical Office, do not contain information on parents.
- Nonetheless, we were able to identify a number of Swiss surveys that can be used for these types of analyses. The results below are based on a selection of these surveys. More surveys are available (especially some older ones) and will be incorporated in future.
- We harmonized the variables in the different surveys to build a common database that can be analyzed in terms of birth cohorts.
- The age range of respondents we analyze is 30 (education) or 35 (class) to 69.

Data: Included Surveys

	Year/Wave	N^a	Label
Un jour en Suisse	1962	29	CH62
Attitudes politiques 1975	1975	390	AP75
Mobilité 81	1981	695	MOB81
ISSP "Social inequality"	1987	412	ISSP87
	1999	968	ISSP99
Les Suisses et leur société	1991	1272	CH91
Swiss Environmental Survey	1994	2220	UWS94
	2007	1973	UWS07
Swiss Labor Market Survey 1998	1998	2328	SAMS98
Swiss Household Panel	1999	5346	SHP99
	2004	2404	SHP04
European Social Survey	2002	1448	ESS02
	2004	1457	ESS04
	2006	1264	ESS06
	2008	1187	ESS08
	2010	984	ESS10
	2012	945	ESS12
MOSAiCH (ISSP)	2005	737	MOS05
	2007	677	MOS07
	2009	845	MOS09
	2011	817	MOS11
	2013	832	MOS13
European Values Study 2008	2008	825	EVS08
Statistics on Income and Living Conditions	2011	6746	SILC11
Total		36801	

[a] Number of observations available for our analyses.

Data: Number of Observations by Birthyear



Data: Classification of Education

Educational level	Included educational degrees
Compulsory or less	No formal education; compulsory education; one year vocational training
Secondary vocational	Vocational training and education; general education without baccalaureate
Secondary general	General education with baccalaureate; vocational baccalaureate; college of education (without university of education)
Tertiary vocational	Professional education and training; advanced federal professional and training diploma; professional education college; university of applied sciences; university of education
Tertiary academic	University; Federal Institute of Technology

Data: Social Class Scheme

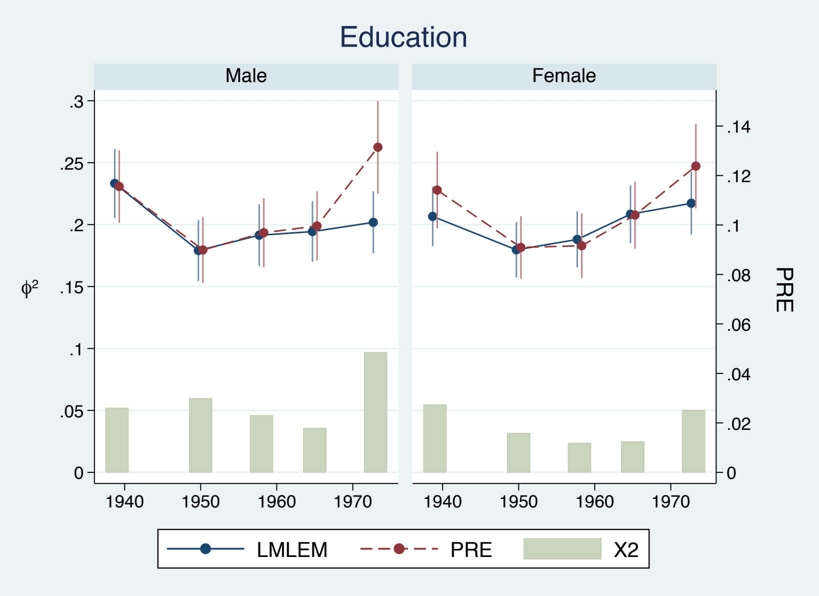
EGP Class		Description
I	Upper service	Higher-grade professionals, administrators and officials; managers in large industrial establishments; large proprietors
II	Lower service	Lower-grade professionals, administrators and officials; higher-grade technicians; managers in small business and industrial establishments; supervisors of non-manual employees
III	Non-manual employees	Routine non-manual employees in administration and commerce; sales personnel; other rank-and-file service workers
IVa,b	Self-employed	Small proprietors, artisans, etc., with employees (IVa); without employees (IVb)
IVc, VIIb	Farmers	Farmers and smallholders, self-employed fishermen (IVc); Agricultural workers (VIIb)
V, VI	Technicians and skilled workers	Lower-grade technicians; supervisors of manual workers; skilled manual workers
VIIa,b	Semi-/unskilled workers	Semi- and unskilled manual workers

Based on Erikson, Goldthorpe and Portocarero (1983: 307).

Results

- Comparison of LMLEM and PRE
- Smoothed PRE
- Extensions
  - PRE with control variables
  - PRE with multiple origin variables
  - Direct and indirect origin effects

Results: Comparison of LMLEM and PRE



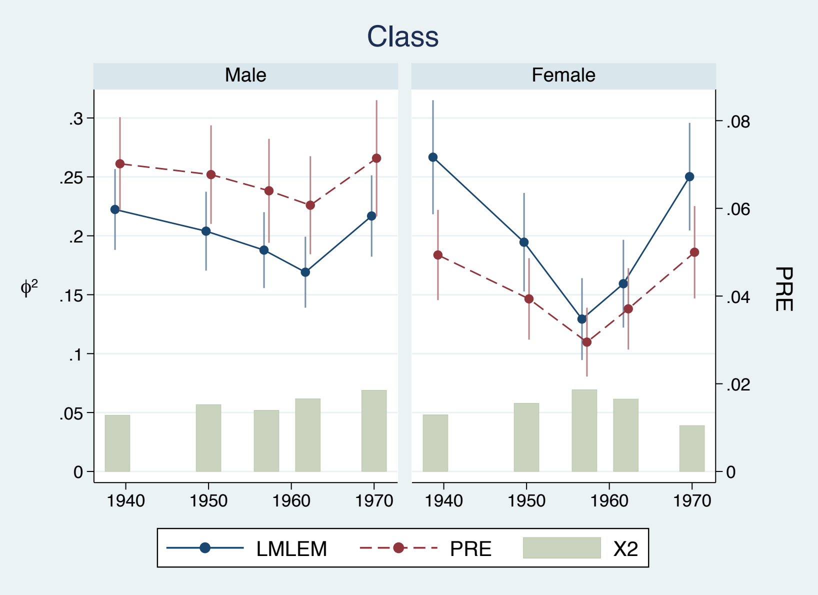
Results: Comparison of LMLEM and PRE

- As discussed above, the LMLEM assumes a common structure of associations between origin and destination categories that remains stable across cohorts.
- Differences between the LMLEM and PRE results may be due to a violation of this assumption.
- We thus included, as grey bars, a cohort-specific goodness-of-fit measure for the LMLEM:

$$\bar{\chi}_k^2 = \frac{1}{N_k} \sum_{i=1}^{I} \sum_{j=1}^{J} \frac{\left(F_{ijk} - \hat{F}_{ijk}\right)^2}{\hat{F}_{ijk}}$$

with F_{ijk} as the observed cell frequencies, $\hat{F}_{ijk}$ as the cell frequencies predicted by the model and N_k as the number of observations in cohort k. High values of $\bar{\chi}_k^2$ indicate bad fit (the scale of $\bar{\chi}_k^2$ is not relevant here and is omitted).

Results: Comparison of LMLEM and PRE



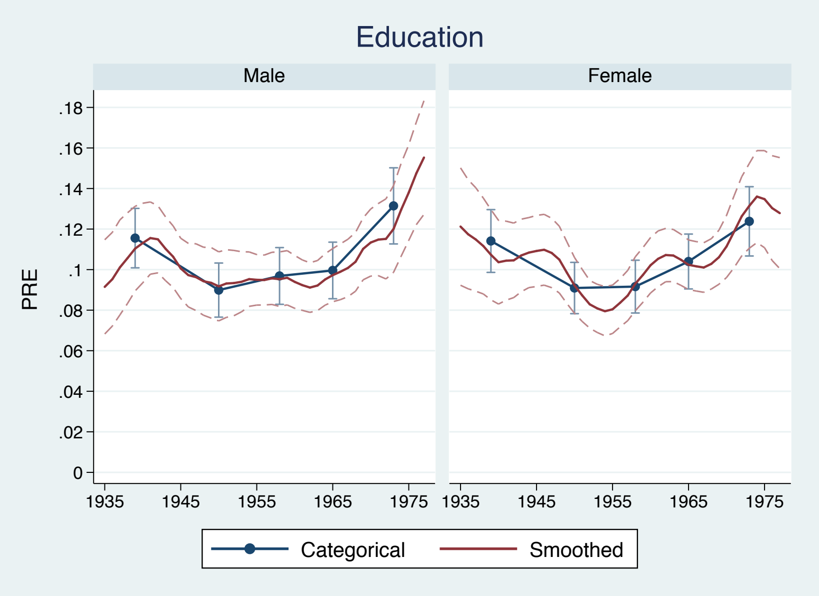
Smoothed PRE

- Instead of dividing the sample in to arbitrary cohorts, PRE can also be computed for each birth year. Results would, however, be very unstable due to the small sample sizes.
- To stabilize results one can include data from surrounding years using kernel weights.
- We use weights

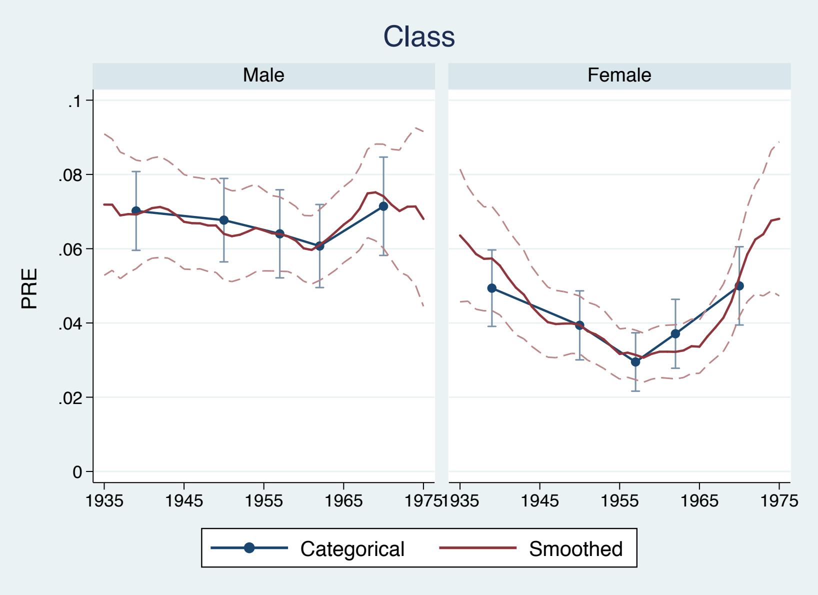
$$w_i(t^*) = w_i \cdot \frac{1}{h} K\left(\frac{t^* - t_i}{h}\right)$$

where t^* is the target birth year, t_i is observations i's birth year, and $K()$ is the Epanechnikov kernel. We set bandwidth h to 5, so that the data window covers a maximum of ± 4 years around target birth year.

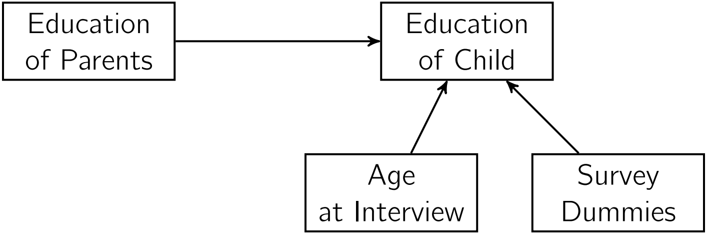
Results: Smoothed PRE



Results: Smoothed PRE



PRE with control variables



PRE with control variables

- Control variables can easily be incorporated into the analysis by including them in both multinomial models.
- PRE then only measures the contribution of parent information over and above these variables.
- Takes into account heterogeneous marginal distribution (e.g., between surveys or age groups).

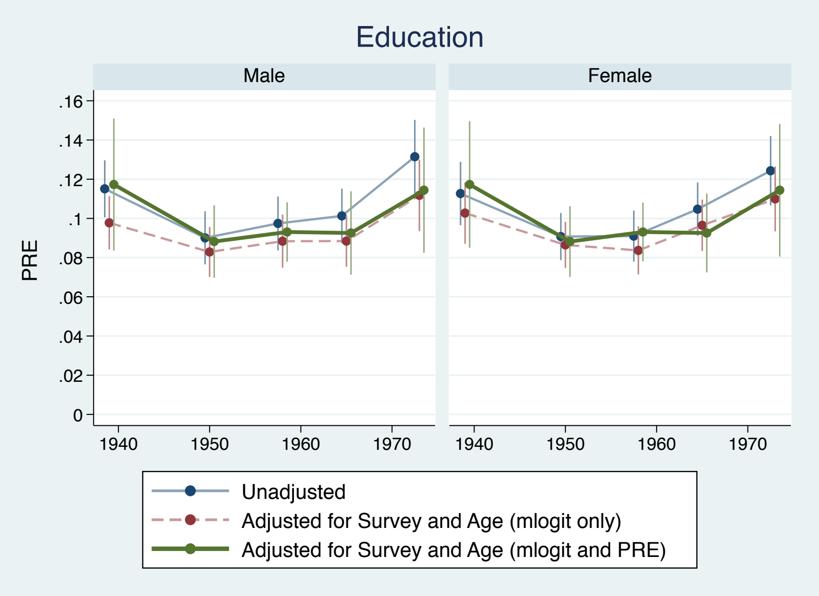
PRE with control variables

- PRE itself may depend on covariates (e.g., PRE will depend on survey when measurement quality differs by survey).
- Therefore, we want to compute the cohort-specific PRE net of the effects of these covariates.
- This can be done as follows:
  - Compute individual-level PRE components:

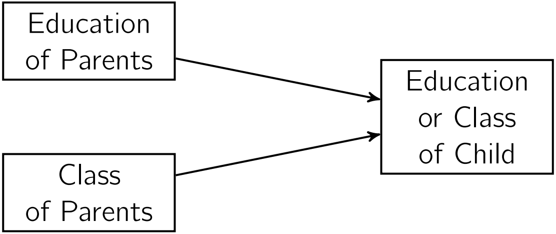
$$\delta_i = \frac{\ln(\widehat{p}_0(Y = y_i)) - \ln(\widehat{p}_1(Y = y_i))}{\widehat{E}_0}$$

  - Regress δ_i on cohort dummies and control variables.
  - Obtain cohort specific PRE values as predictive margins.

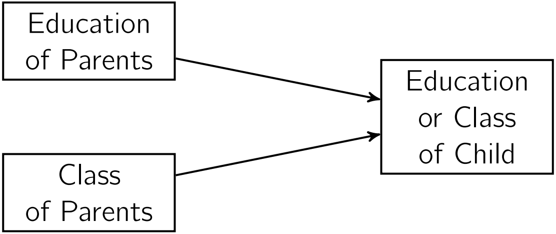
Results: PRE with control variables



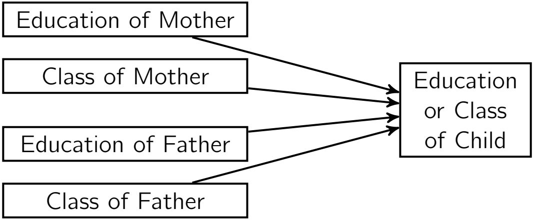
PRE with multiple origin variables



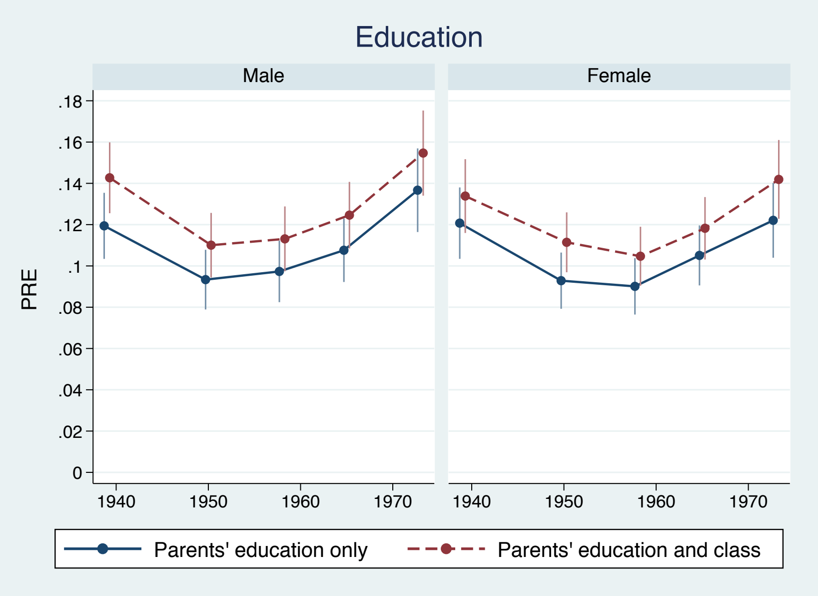
PRE with multiple origin variables



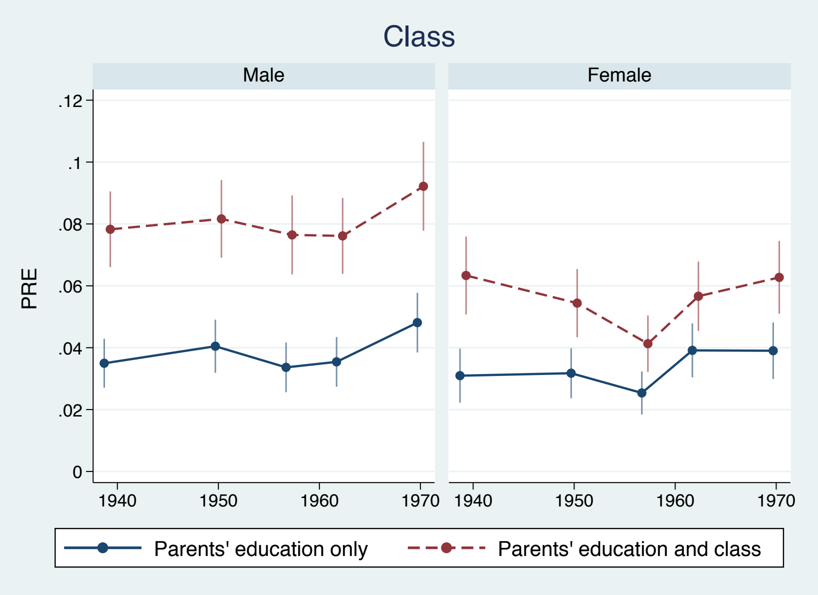
Could also get more complicated such as:



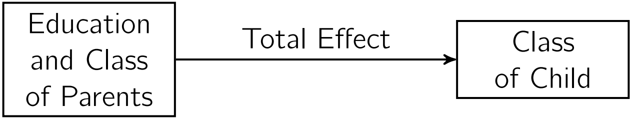
Results: PRE with multiple origin variables



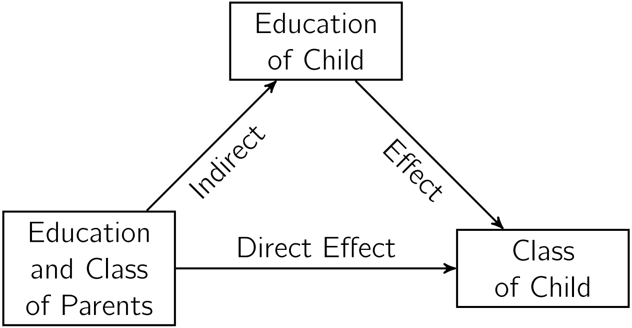
Results: PRE with multiple origin variables



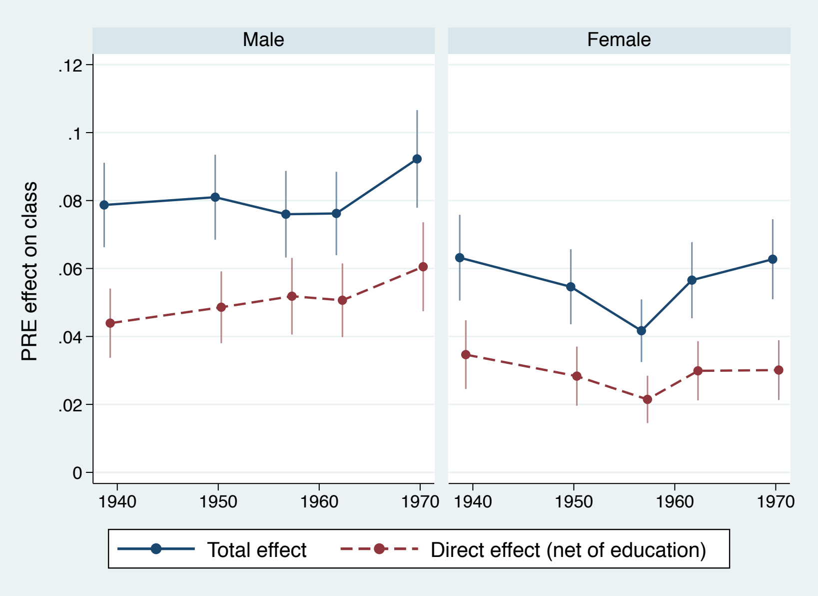
Direct and indirect origin effects



Direct and indirect origin effects



Results: Direct and indirect origin effects



Conclusions

- The PRE approach seems to be a viable and flexible model to analyze social mobility.
  - It produces results that are comparable to the classic LMLEM. Deviations between PRE and LMLEM seem to be related to misfit of the LMLEM.
  - It can easily include multiple origin variables and control variables.
    - Likewise: Analysis of the effects of context variables on PRE to explain changes/differences in effects of social origin.
  - It has a clear interpretation (proportional reduction of prediction errors): How much does the knowledge of parents' characteristics improve the prediction of the child's status?

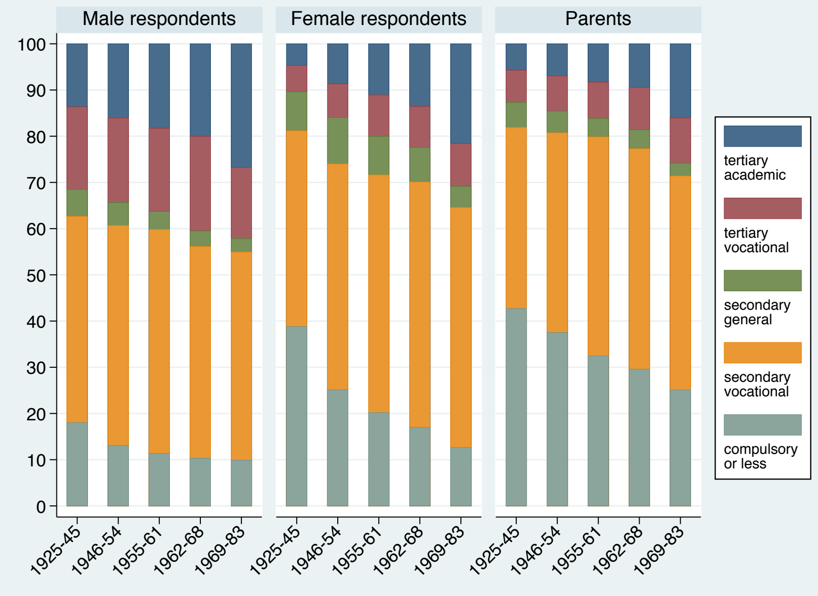
Conclusions

- Substantive conclusions
  - Our results indicate that social mobility increased among birth cohorts in the mid 1930s to about 1960, but then started to decrease again.
  - In general, this pattern can be observed for both men and women and both education and class. The pattern, however, is least pronounced for men's class.
  - For respondent's education the PRE approach leads to more pronounced results than LMLEM. This indicates that the structure of association changed over time for education.
  - Net of parents education, parents' class still has an effect on both respondent's education and class. As expected, the effect on class is stronger.
  - Parents characteristics have a clear direct effect on respondent's class, net of respondent's education.

References

- Erikson, Robert, John H. Goldhorpe. 1992. The constant flux. A study of class mobility in industrial societies. Oxford: Clarendon.
- Erikson, Robert, John H. Goldthorpe, Lucienne Portocarero (1983). Intergenerational Class Mobility and the Convergence Thesis: England, France and Sweden. The British Journal of Sociology 34(3): 303–343.
- Jann, Ben, Benita Combet. 2012. Zur Entwicklung der intergenerationalen Mobilität in der Schweiz. Swiss Journal of Sociology 38(2): 177–199.
- Theil, Henri. 1970. On the estimation of relationships involving qualitative variables. American Journal of Sociology 76(1): 103–154.
- Xie, Yu (1992). The Log-Multiplicative Layer Effect Model for Comparing Mobility Tables. American Sociological Review 57(3): 380–395.

Data: Education by Birth Cohort



Data: Social Class by Birth Cohort